Online Database Coverage of Publications by Biology Faculty at a Research University: A Comparative Study

A Paper on the General Topic of Applied Library Research for Presentation at the 1989 Annual Conference of the California Library Association

Daniel Taylor
Science Librarian
California Institute of Technology

September 1989

Online Database Coverage of Publications by Biology Faculty at a Research University: A Comparative Study

Making the optimal choice of an online database or databases to search for current literature in a particular research field may not be as simple as it seems. BIOSIS may come to mind, for example, as the logical first choice for retrieving current publications in biological research, to be supplemented by other databases if exhaustive retrieval is required. This may not be the case, however, if initial searching of databases less expensive than BIOSIS can retrieve the relevant literature almost as effectively. If so, searching BIOSIS may be unnecessary or inappropriate, depending upon search objectives and priorities.

If retrieval of conference literature is important, BIOSIS should be searched because its Biological Abstracts/RRM segment may include conference papers that other databases omit. The costs of searching BIOSIS, when needed, may be minimized, moreover, by incurring "citation charges" only for those titles not already retrieved from the less expensive databases.

Coverage of many aspects of biological research by Chemical Abstracts is quite extensive, and for those eligible for STN's academic discount (approximately 70%), retrieval from CA File is available at very low cost. Similarly coverage of biological research by MEDLINE is also

extensive, and retrieval is very economical for those who access MEDLINE directly via NLM's MEDLARS.

The present study quantifies the collective retrieval from these databases in comparison to retrieval from BIOSIS. This is accomplished by searching all three databases for the recent publications of the biology faculty at a research university having typical emphases at the cellular and molecular levels and including genetics, neurobiology, and psychobiology.

Methodology

Author searches on the 29 Biology Division professorial faculty at the California Institute of Technology were performed in CA File, MEDLINE, AND BIOSIS during July 1989. The searches in each file were extended back through the online equivalent of the January 1987 issues of Chemical Abstracts, Index Medicus, and Biological Abstracts; functions used to do this were ranging, stringsearching, and limiting, respectively. Expand or truncation capabilities were always utilized to retrieve all variant forms of the name in the database. Bibliographic citations were printed from CA File and MEDLINE; DIALOG's format 6 was used to print from BIOSIS in most cases. All authors were searched in BIOSIS, then all in CA File, and finally all in MEDLINE.

Printouts were then grouped by individual author. For a particular author, CA File citations were examined first since the institutional affiliation facilitated verification of proper authorship in most cases; actual articles were checked if necessary to make a final determination. Verified CA File citations were then checked against and noted on the BIOSIS printouts, except in the cases where no match was found. (Examples are given in Appendix A.) MEDLINE citations were then checked against and noted on the BIOSIS printouts, except where no match was found. (Examples: Appendix B.) The BIOSIS numbers were used to look up the print citations to verify proper authorship as needed for BIOSIS titles. Actual articles were checked in the cases deemed necessary for either the MEDLINE printouts or the BIOSIS print citations.

Verified citations were then counted and tabulated for each author for each database. On each BIOSIS printout, the number of Biological Abstracts (as opposed to Biological Abstracts/RRM) citations was also counted and tabulated; the differentiation was apparent because the BIOSIS numbers for the two print publications are in separate ranges. (Examples of citations retrieved from both segments of BIOSIS, but not from CA File nor MEDLINE, are given in Appendices C and D.)

On each BIOSIS printout the combined total of CA File and MEDLINE citations appearing in Biological Abstracts was determined and calculated as a percentage; so also the

combined total of CA File and MEDLINE citations appearing in BIOSIS overall, also calculated as a percentage. Individual counts were necessary in each case, because both CA File and MEDLINE retrieved some citations not retrieved by BIOSIS.

Collective total verified retrievals were determined for CA File, MEDLINE, BIOSIS, and the Biological Abstracts segment of BIOSIS. Collective total CA File plus MEDLINE retrievals as percentages of BIOSIS and of its Biological Abstracts segment were calculated.

Results

For the 29 biology faculty authors searched, CA File gave 274 verified retrievals, MEDLINE gave 266, and BIOSIS gave 570, of which 235 were in the Biological Abstracts segment.

CA File plus MEDLINE retrieval was 92.8 % of the Biological Abstracts retrieval. CA File plus MEDLINE retrieval was 53.7 % of the total BIOSIS retrieval.

For 20 of the 29 faculty, CA File plus MEDLINE retrieved 100 % of the Biological Abstracts retrieval. 9.5 % of the CA File citations, and 7.1 % of the MEDLINE citations, were not retrieved by BIOSIS.

Conclusion

Online searches of CA File, MEDLINE, and BIOSIS for the recent publications of a typical biology research faculty demonstrate that the combined journal citation retrieval from CA File and MEDLINE compares very favorably with retrieval from BIOSIS. In many cases the CA File and MEDLINE retrievals may be sufficient, depending upon one's search objectives and priorities. In those instances where the conference literature or exhaustive retrieval are important, as many as 50 - 90 % of the BIOSIS citation charges may be avoided by more economical initial searches of CA File and MEDLINE. Both CA File and MEDLINE may also retrieve significant quantities of relevant citations not picked up by BIOSIS. Therefore BIOSIS may often not be the logical first choice for retrieving current publications in research biology.

APPENDIX A

EXAMPLES OF CITATIONS RETRIEVED FROM CA FILE BUT NOT BIOSIS

CA107(25):231501h Sequence diversity of gap junction proteins. Revel, J. P.; Yancey, S. B.; Nicholson, B.; Hoh, J. (Div. Biol., California Inst. Technol., Pasadena, CA 91125, USA). Ciba Found. Symp., 125(Junctional Complexes Epithelial Cells), 108-27 (Eng) 1987. CODEN: CIBSB4. ISSN: 0300-5208.

CA108(5):34627y Spliceosome assembly in yeast. Cheng, Soo Chen; Abelson, John (Div. Biol., California Inst. Technol., Pasadena, CA 91125, USA). Genes Dev., 1(9), 1014-27 (Eng) 1987. CODEN: GEDEEP. ISSN: 0890-9369.

APPENDIX B

EXAMPLES OF CITATIONS RETRIEVED FROM MEDLINE BUT NOT BIOSIS

```
prt compr

PROG:

1
UI  - 89100779
AU  - Grzywacz NM ; Koch C
TI  - Functional properties of models for direction selectivity in the
      retina.
SO  - Synapse 1987;1(5):417-34

SS 2 /C?
USER:
prt compr

PROG:

1
UI  - 89163913
AU  - Hunkapiller T ; Hood L
TI  - Diversity of the immunoglobulin gene superfamily.
RF  - REVIEW ARTICLE: 266 REFS.
SO  - Adv Immunol 1989;44:1-63

SS 3 /C?
USER:
```

APPENDIX C

EXAMPLES OF CITATIONS RETRIEVED FROM BIOSIS (BIOLOGICAL ABSTRACTS SEGMENT) BUT NOT FROM CA FILE NOR MEDLINE

7/3/1
0017744368 BIOSIS Number: 84116653
BENDING PATTERNS OF CHLAMYDOMONAS FLAGELLA IV. MUTANTS WITH DEFECTS IN INNER AND OUTER DYNEIN ARMS INDICATE DIFFERENCES IN DYNEIN ARM FUNCTION
BROKAW C J; KAMIYA R
DIV. BIOL., CALIF. INST. TECHNOL., PASADENA, CALIF. 91125.
CELL MOTIL CYTOSKELETON 8 (1). 1987. 68-75. CODEN: CMCYE
Language: ENGLISH
?

8/3/1
0019559558 BIOSIS Number: 88026476
GENES DIRECTING FLOWER DEVELOPMENT IN ARABIDOPSIS
BOWMAN J L; SMYTH D R; MEYEROWITZ E M
DIV. BIOL., 156-29, CALIF. INST. TECHNOL., PASADENA, CA 91125.
PLANT CELL 1 (1). 1989. 37-52. CODEN: PLCEE
Language: ENGLISH
?

APPENDIX D

EXAMPLES OF CITATIONS RETRIEVED FROM BIOSIS (BIOLOGICAL ABSTRACTS/RRM SEGMENT) BUT NOT FROM CA FILE NOR MEDLINE

9/3/1
0019010645 BIOSIS Number: 36010645
SUBCELLULAR LOCALIZATION OF MESSENGER RNA SPECIES IN DROSOPHILA RETINA REVEALED BY LIGHT AND ELECTRON MICROSCOPIC IN SITU HYBRIDIZATION
POLLOCK J A; BENZER S; DEERINCK T; ELLISMAN M H
DIV. BIOL. 156-29, CAL. INST. TECHNOL., UCSD, SAN DIEGO, CA.
18TH ANNUAL MEETING OF THE SOCIETY FOR NEUROSCIENCE, TORONTO, ONTARIO, CANADA, NOVEMBER 13-18, 1988. SOC NEUROSCI ABSTR 14 (1). 1988. 732.
CODEN: ASNEE
Language: ENGLISH
?

10/3/1
0019525338 BIOSIS Number: 37014253
BACTERIAL DNA INVERSION SYSTEMS
GLASGOW A C; HUGHES K T; SIMON M I
DEP. OF BIOL. 147-75, CALIF. INST. OF TECHNOL., PASADENA, CALIF. 91125.
BERG, D. E. AND M. M. HOWE (ED.). MOBILE DNA. XVII+972P. AMERICAN SOCIETY FOR MICROBIOLOGY: WASHINGTON, D.C., USA. ILLUS. MAPS. ISBN 1-55581-005-5.
0 (0). 1989. 637-660. CODEN: 40261
Language: ENGLISH
?